Unsupervised Learning (Clustering) of Odontocete Echolocation Clicks

Marie A. Roch
Dept. of Computer Science
San Diego State University
5500 Campanile Drive
San Diego, CA 92182-7720
phone: (619) 594-5830 fax: (619) 594-6746 e-mail: marie.roch@sdsu.edu

Margareta Ackerman
Dept. of Computer Science
Florida State University
253 Love Bldg.
Tallahassee FL, 32306
phone: (850) 645-4029 fax: (850) 644-0058 e-mail: mackerman@fsu.edu

Simone Baumann-Pickering
Scripps Institution of Oceanography
University of California at San Diego
9500 Gilman Dr.
La Jolla, CA 92093-0205
phone: (858) 534-7280 fax: (858) 534-6849 e-mail: sbaumann@ucsd.edu

Award Number: N00014-15-1-2299
http://roch.sdsu.edu

LONG-TERM GOALS

The primary long-term goal is to develop methods for clustering of marine mammal echolocation clicks to learn about species assemblages where little or no prior knowledge exists about these species' signal repertoire. Being able to monitor individual species can be effective, for understanding how certain species use a region and habitat or how naval operations may affect their behavior.

OBJECTIVES

In the first phase of this project, we will work on data from the Southern California Bight where many of the species can be acoustically identified, enabling the development of clustering algorithms whose performance can be evaluated. The option year of this project will look at areas where our understanding is less developed such as data collected from the Gulf of Mexico or the Atlanic.

APPROACH

Acoustic encounters with odontocetes are detected automatically and noise-corrected cepstral features representing the click spectrum are collected for each echolocation click. A probabilistic model of the distribution of each encounter is created. A symmetric Kullback-Leibler (KL) divergence metric is computed between every pair of distributions through the use of a Monte Carlo simulation, followed by the application of an unsupervised learning algorithm.

WORK COMPLETED

In the first several months of this project, distributions were estimated for every detected encounter in the development data of the 2015 International Workshop on Detection, Classification, Localization, and Density Estmation of Marine Mammals Using Passive Acoustic Monitoring (DCLDE). KL divergence maps were created for all known species, but the sperm whale (*Physeter macrorhynchus*) and porpoise (*Phocoenidae spp.*) data were rejected from the preliminary analysis as the feature set being used did not cover enough of their bandwidth. Results on Risso's dolphins (*Grampus griseus*), Pacific white-sided dolphins (*Lagenorhynchus obliquidens*), and Cuvier's beaked whales (*Ziphius cavirostris*) showed good potential for clustering.

RESULTS

In our analysis of clusters, we were able to examine pairs of acoustic encounters to see which ones appeared to be closely related to one another. We noted that some of the Pacific white-sided and Risso's dolphin acoustic encounters were closer to one another than we would have expected. This led us to reexamine our echolocation click detection and feature extraction process and identify a source of false detections triggered by periodic data logger self-noise events. We have developed a method to correct for this and are seeing marked improvements not only in this project, but other related classifiers. This provides a new tool for evaluating the efficacy of DSP feature extraction streams.

The distance map between the three species that we have processed so far shows good separation and should be clusterable. The group of odontocetes that we cannot label reliably by their acoustic features, primarily common dolphins (*Delphinus spp.*) and bottlenose dolphins (*Tursiops truncatus*), is somewhat separated from these species, but some of these encounters appear somewhat close to the above three species and we are investigating the cause of this. A preliminary version of these results was presented at the 2015 DCLDE workshop.

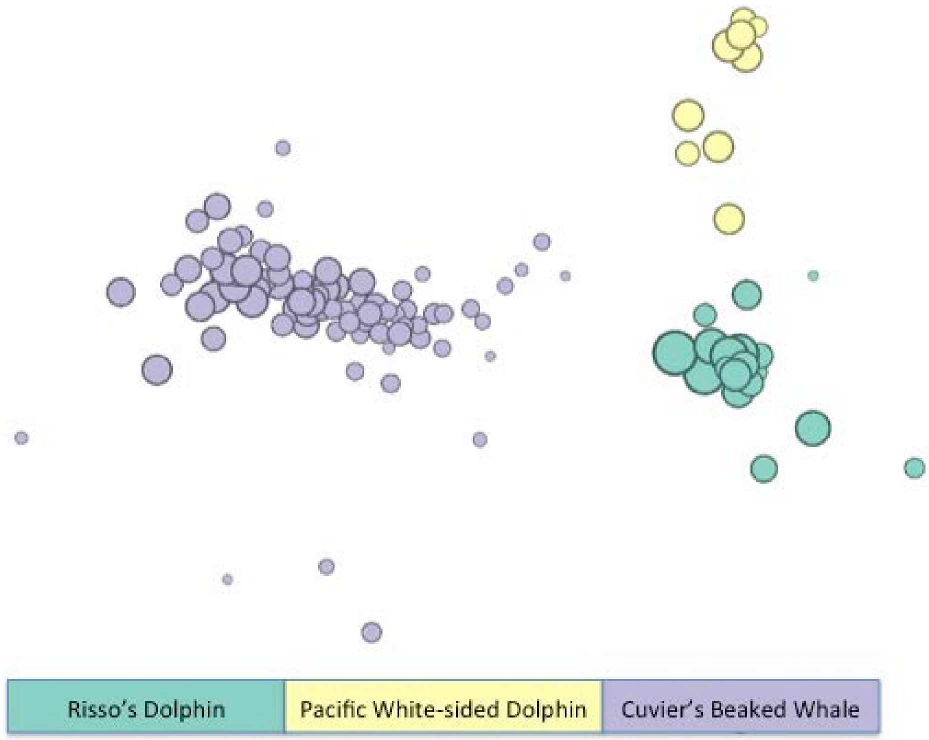

Figure 1 – Separation of Pacific white-sided dolphin (yellow), Risso's dolphin (green), and Cuvier's beaked whale (purple) acoustic encounters based on KL divergence. Each circle represents an acoustic encounter; circle size is representative of the number of echolocation clicks detected in the encounter with the smallest acoustic encounters having 800 clicks.

IMPACT/APPLICATIONS

Increasingly, the US Navy is depending upon passive acoustic monitoring to mitigate impact on cetaceans during naval operations. For some of these areas, marine mammal species presence and at least some of the species' acoustic repertoire has been documented (e.g. The Southern California Offshore Range, The Atlantic Undersea Test and Evaluation Center, and The Hawai'i Range Complex). Other areas of naval use have a larger degree of uncertainty. As the Navy increases the use of acoustic monitoring, robust classification techniques will continue to gain importance and preference may be given to techniques that can be implemented in areas where species acoustic repertoire is less well known or where species-specific signals are difficult to discriminate and classify. This will provide the US Navy with new tools for informative, signal type- or species-level acoustic monitoring and successful mitigation measures to reduce takes of marine mammals during naval operations.

RELATED PROJECTS

N39430-15-C-1712 – Tethys, a workbench for passive acoustic monitoring metadata. PI Marie Roch, Simone Baumann-Pickering, Ana Širović. Database used for managing acoustic metadata providing automated access to instrument characteristics.

ONR N00014-13-IP20051– Advanced Methods for Passive Acoustic Detection, Classification, and Localization of Marine Mammals. PI Jonathan Klay, Dave Mellinger, Dave Moretti, Steve Martin and Marie A. Roch. Techniques developed in this project form the basis of the detection signal processing chain used in the current project.